Methodological guidelines for developing accident modification functions

Rune Elvik [a] [b]

a Institute of Transport Economics, Gaustadalleen 21, NO-0349 Oslo, Norway

b Department of Civil Engineering, Aalborg University, Sofiendalsvej 11, DK-9210 Aalborg SV, Denmark

ABSTRACT

This paper proposes methodological guidelines for developing accident modification functions. An accident modification function is a mathematical function describing systematic variation in the effects of road safety measures. The paper describes ten guidelines. An example is given of how to use the guidelines. The importance of exploratory analysis and an iterative approach in developing accident modification functions is stressed. The example shows that strict compliance with all the guidelines may be difficult, but represents a level of stringency that should be strived for. Currently the main limitations in developing accident modification functions are the small number of good evaluation studies and the often huge variation in

estimates of effect. It is therefore still not possible to develop accident modification functions for very many road safety measures.

Key words: accident modification function; methods; guidelines; road safety; evaluation studies

1 INTRODUCTION

There is a growing understanding of the fact that the effects of road safety measures vary systematically (Hauer et al. 2012). It is therefore not always very informative to state these effects in terms of a single point estimate. An accident modification function can provide a more informative and precise description of effects, by statistically modelling variation in effects as a function of one or more independent variables.

Developing accident modification functions is, however, not easy and requires careful attention to the quality of evaluation studies and to whether the distribution of estimates of effect in these studies displays a systematic pattern. The objective of this paper is to propose methodological guidelines for developing accident modification functions. The guidelines address the following questions:

1. How should studies serving as the basis for developing an accident modification function be selected?

2. What types of preparatory analyses are required before starting to develop an accident modification function?

3. How can independent variables in an accident modification function be identified?

4. How can outlying data points be identified?

5. How can the most suitable mathematical form of an accident modification function be determined?

6. How can one decide whether a single or more than one accident modification function best fits the data?

7. How can the quality of an accident modification function be evaluated?

8. How can the effects of analytic choices made when developing an accident modification function be evaluated (in terms of sensitivity analysis)?

9. How can heteroscedastic data best be analysed when developing an accident modification function?

10. How can accident modification functions be updated?

Ten guidelines addressing these issues are proposed. Each guideline is illustrated by an example showing how to use the guideline. All examples refer to studies of the effects on accidents of speed enforcement. The guidelines proposed are listed in Table 1. In the following sections, each guideline will be presented in detail.

Table 1 about here

2 CLASSIFY, CODE AND SELECT STUDIES

The first step in developing an accident modification function is to identify the studies that will serve as a basis for developing the function. A systematic literature survey should be made to identify relevant studies. Once relevant studies have been identified, they should be classified according to study design and how well they control for potentially confounding factors. This is an essential preparatory step for analysis, because studies employing different designs do not control for the same potentially confounding factors and tend to produce different estimates of effect.

Table 2 proposes a classification of study designs and identifies three levels of study quality for each design. Four common types of study design are listed in Table 2. For

each study design, studies are classified as high, medium or low quality depending on how well they control for potentially confounding factors.

Table 2 about here

Whenever possible, one should avoid mixing studies using different designs, or not controlling for the same confounding factors, when developing an accident modification function. If there are enough studies to discard those of medium or low quality, doing so is recommended. If most studies are of medium or low quality, one should not try to develop an accident modification function, as there is a non-negligible risk that it will be biased and misleading.

An example of the selection of studies for use in developing an accident modification function is given in Table 3. The studies listed in Table 3 were, except for the most recent one, retrieved in an earlier study (Elvik 2011) that developed an accident modification function for speed enforcement.

Table 3 about here

Studies were only included if the following criteria were fulfilled:

1. The study employed an experimental or observational before-after design.
2. The dependent variable was accidents.
3. For studies containing multiple estimates of effect, these should display an internally systematic pattern (see below).
4. For studies containing a single estimate of effect, this should be consistent with theoretical predictions (see below).

The pattern of results in a study is internally consistent if it shows, allowing for random variation, a dose-response pattern. This means that increasing enforcement is associated with an accident reduction; reducing enforcement is associated with an increase of accidents. If only a single estimate of effect is provided, it is consistent with theoretical predictions if it shows that reducing enforcement is associated with an increase in the number of accidents and increasing enforcement is associated with a reduction of accidents. It might seem dubious to omit studies when their results are not, at least broadly, consistent with a theoretically expected pattern. It looks like omitting results we do not like. However, one might just as well argue that road safety evaluation research is too rarely guided by theory; that too few results are ruled out as theoretically implausible.

A sensitivity analysis was made by including the studies that were omitted from the main analysis. For each study, the following independent variables were coded:

1. Country where study was made

2. Publication year

3. Levels of enforcement studied

4. Whether speed cameras were used or not

Table 4 shows the final data table. There were 31 estimates of effect in total. The principal independent variable of interest is the level of enforcement. For this variable, 1 is the baseline level, 0 is no enforcement and 2 is twice the baseline level.

Table 4 about here

3 PREPARATORY ANALYSIS

Three issues should be addressed in the preparatory analysis. First, the contributions of random and systematic variation to the total variation in study findings should be determined. The I^2 statistic in meta-analysis is a useful indicator of the relative contribution of systematic variation to the overall variation in estimates of effect (Borenstein et al. 2009). It is best stated as a percentage and should, as a guideline, have a value greater than 50 percent.

Second, the distribution of estimates of effect should be tested for the possible presence of publication bias. Publication bias denotes a tendency not to publish findings that are regarded as difficult to interpret, or unwanted, like finding an increase in the number of accidents when the opposite was expected. All tests for publication bias are based on assumptions that cannot be tested directly (Rothstein et al. 2005). Nevertheless, if publication bias is indicated, an accident modification function should not be developed.

Third, the effects of country and publication year on estimates of effect should be examined. Country and publication year are basically confounding variables, since the ambition of research is to develop knowledge that is internationally transferable and represents comparatively stable relationships.

For the studies of speed enforcement listed in Table 4, the I^2 statistic had a value of 97.7, indicating that almost all variation in estimates of effect is systematic. The value of τ^2 was 0.061. Thus, there is clearly a substantial systematic variation in estimates of effect. To test for publication bias, a funnel plot was developed and analysed by means of the trim-and-fill method (Duval 2005). There was no indication of

publication bias. The funnel plot indicated that even estimates of effect based on small standard errors varied considerably. It is reproduced in Figure 1.

Figure 1 about here

To test for the effects of country and publication year, meta-regressions were run using a macro for SPSS developed by Lipsey and Wilson (2001). Two runs were made, each time omitting one country; otherwise the dummies identifying countries will be perfectly collinear, leaving no degrees of freedom to fit model coefficients. Greece was omitted in the first run, the United States in the second. None of the country variables had statistically significant coefficients in the first run. In the second run, the coefficient for Australia was statistically significant, which it was very far from being in the first run. Publication year was not statistically significant in either run. The dummy variable identifying Australia is perfectly collinear with the dummy for use of a speed camera. The meta-regressions run to test the effects of country and publication year did not include the speed camera indicator. The coefficient estimated for Australia will include the effect of the speed camera dummy. It is concluded that neither country of origin nor publication year are likely to produce residual confounding in a model not including these variables.

4 IDENTIFYING INDEPENDENT VARIABLES

The independent variables of primary interest are of two types:

1. Characteristics of the road safety measure, such as indicators of its quality or standard or extent of use.

2. Characteristics of the context for use of the road safety measure, such as the type of traffic environment it is used in.

In the study of speed enforcement used to illustrate the guidelines, the following independent variables were used:

1. Level of enforcement, a numerical variable stated with one decimal and ranging from 0.0 to 14.0. 1.0 is the baseline (current) level of enforcement.
2. Use of speed camera, which is dummy variable taking the value of 1 if a speed camera was used. The extent of use is stated as camera hours and is included in the level of enforcement variable.

5 IDENTIFYING OUTLYING DATA POINTS

An outlying data point is a single data point that has a decisive influence on the summary estimate of effect in meta-analysis or on the functional form of an accident modification function. To identify outlying data points when developing an accident modification function, the following procedure is recommended:

1. Develop an initial accident modification function including all data points. The mathematical form of this function may be subject to revision at subsequent stages of analysis.
2. Examine a cumulative residuals plot for the initial function in order to identify sudden jumps that may indicate the presence of outlying data points.

3. Inspect a plot of the function and estimate standardised residuals. On the average, if 95 % confidence limits are applied, about 1 in 20 residuals should be above or below two standard errors from the fitted function.

The following initial accident modification function was fitted to the data listed in Table 4:

$$Ln(\text{estimate of effect}) = \alpha + \beta_1 level + \beta_2 cameras$$

A maximum likelihood meta-regression was run. The coefficient estimates were 0.1490 for the constant term (α), –0.0581 for the level term (β_1) and –0.2358 for the camera dummy (β_2). All coefficients were highly statistically significant. A cumulative residuals plot (Hauer and Bamfo 1997, Hauer 2015) was developed it is shown in Figure 2.

Figure 2 about here

The cumulative residuals are well-behaved until the level of enforcement reaches the value of 9. The residuals then take a sudden jump up and go outside the 95 % confidence limits indicated by the dotted curves. This suggests the presence of an outlying data point.

The fitted function is shown in Figure 3. It consists of two parts. The upper curve is the function fitted to data points that did not involve the use of speed cameras (i.e. the speed camera dummy had a value of 0). The lower curve is fitted to data points involving the use of speed cameras. Since the two curves differ only by the inclusion of an additional coefficient in the lower curve, they run in parallel. An examination of

the standardised residuals is very informative. Standardised residuals were estimated as follows:

$$\text{Standardised residual} = \frac{(E_i - M_i)}{\left(\frac{1}{\sqrt{W_i}}\right)}$$

E_i is the i-th estimate of effect (i = 1, 2, …, 31). M_i is the corresponding model estimate of effect according to the preliminary accident modification function. W_i is the fixed-effects statistical weight of the i-th estimate of effect.

Figure 3 about here

The function fits very poorly to the data points involving the use of speed cameras. There are ten of these data points. In terms of level of enforcement (camera hours) they range from 0.7 to 7.1. For all data points from 0.7 to 1.3 (five data points) the standardised residuals are highly positive ranging from 3.17 to 5.72. For all data points from 2.3 to 7.1 (five data points) the standardised residuals are negative, ranging from –0.26 to –7.18. This suggests that separate functions should be fitted to the data points with and without speed cameras, and that these functions need not have the same mathematical form.

Turning to the function fitted to the data points not involving the use of cameras, estimation of standardised residuals confirms that the data point located at level 9 for enforcement is indeed outlying (standardised residual 2.93). The data point located at level of enforcement 3.5 looks suspicious, but is strictly speaking not outlying. It is nevertheless located so far from the other data points that including it when developing an accident modification function hardly adds information of any value. The function fitted to the non-camera data points in figure 3 passes below most data

points located to the right of about 6 for the level of enforcement, suggesting that a function with a stronger curvature would fit the data better.

6 IDENTIFYING THE BEST FITTING FUNCTIONAL FORM – CONVENTIONAL SPEED ENFORCEMENT

The analysis of Figure 3 suggested that the initial accident modification function could be improved, both with respect to the data points not involving the use of speed cameras and with respect to the data points involving the use of speed cameras. Some commonly used functions can be tested on an Excel spreadsheet, in particular linear, exponential, logarithmic, power and polynomial. By testing these functions, one may gain an impression of whether one of them is clearly superior.

Tests made in Excel for the data points not involving the use of speed cameras, including the two data points labelled as outlying in Figure 3, indicated that neither a linear, an exponential nor a polynomial function were clearly to be preferred. All these functions fitted the data quite poorly. The logarithmic and power functions could not be fitted as one of the data points had the value of zero. It was decided to:

1. Omit the two data points identified as outlying in Figure 3.
2. Define a new variable, level of enforcement squared.

A function was then fitted using ln(estimate of effect) as dependent variable and level of enforcement and level of enforcement squared as independent variables. Inspection of a cumulative residuals plot indicated that residuals became highly negative at a value of about 1 for level of enforcement (i.e. the current level) and

strayed outside the 95 % confidence limits of the plot. To try to remedy this problem, three data points, all referring to the current level of enforcement (level 1.0) were merged into a single data point. The three data points were all from the same study (Shoup 1973). The estimate of effect for the merged data point was 0.986, which is close to the theoretically expected value of 1.0. The function was the fitted again.

The cumulative residuals plot improved a little, but did still not look ideal. However, as other criteria of model quality are used in addition to the cumulative residuals plot, the function was provisionally accepted. Its quality is assessed more systematically in section 8 of the paper. Figure 4 shows the function and the data points to which it was fitted.

Figure 4 about here

7 IDENTIFYING THE BEST FITTING FUNCTIONAL FORM – USE OF SPEED CAMERAS

As far as the data points referring to the use of speed cameras are concerned, tests in Excel, see Figure 3, indicated that an exponential function would fit the data quite well. An exponential accident modification function was therefore fitted to the data. Figure 5 shows this function and the data points it was fitted to.

Figure 5 about here

The function appears to fit the data quite well. A more formal assessment is reported in the following section.

8 ASSESSING MODEL QUALITY

The following criteria are proposed regarding the quality of accident modification functions:

1. Overall goodness-of-fit assessed in terms of a cumulative residuals plot and the value of the residual systematic variation (τ^2).

2. Unbiasedness of model predictions: the model should not, on the average predict a larger or smaller effect than the data points serving as its basis

3. Normality in the distribution of standardised residuals

4. Heteroscedasticity in the standardised residuals

5. Autocorrelation of residuals

These statistics are reported in Table 5 for the two accident modification functions developed for speed enforcement.

Table 5 about here

In both data sets, systematic variation in estimates of effect contributed to nearly all the variation (more than 90 percent). The accident modification functions fitted were able to explain more than 90 percent of the systematic variation in estimates of effect.

The cumulative residuals plots were not ideal for any of the two functions, but did at least contain more than 80 percent of the data points inside the confidence limits. A visual examination of Figures 4 and 5 does not suggest that the functions fit the data poorly. Both models were found to be unbiased when model coefficients were used

to predict the mean effect, i.e. the fitted functions do not systematically predict too many or too few accidents. Standardised residuals were normally distributed for the function fitted to conventional speed enforcement, but more widely dispersed than normal for the function fitted to the use of speed cameras. All the data points for speed cameras had large statistical weights and therefore small standard errors.

A potentially serious problem when developing accident modification functions is the heteroscedasticity of the data. This means that not all data points have the same sampling variance. Data points based on a low number of accidents will have a greater variance than data points based on a high number of accidents. This is an inherent characteristic of nearly all data sets used in meta-analysis of road safety evaluation studies and therefore likely to be a problem when developing accident modification functions. To test for heteroscedasticity of residuals, a graphical method has been applied. The logic of this test can be explained by reference to Figure 6.

Figure 6 about here

Figure 6 plots the standardised residuals for the accident modification function fitted to studies evaluating the effects of conventional speed enforcement. Separate trend lines have been fitted to the positive and negative residuals. If these lines are both horizontal, there is no heteroscedasticity. If, as in Figure 6, the lines converge to the right, there is what might be termed "negative heteroscedasticity", meaning that the differences between the residuals become smaller as a function of the independent variable. Positive heteroscedasticity refers to the opposite situation; that the trend lines move apart. A T-test was applied to determine if the difference in slopes between the trend lines fitted to the positive and negative residuals was statistically

significant. First, the difference in slopes was computed. Then, the standard error of this difference was computed. The value of T was estimated as difference divided by standard error. The degrees of freedom was equal to the minimum number of data points minus 2. In Figure 6 there were eight positive data points; thus the T-test had six degrees of freedom. As can be seen from Table 5, the residuals for the accident modification function for conventional speed enforcement bordered on statistical significance for heteroscedasticity. The residuals for the speed camera function were clearly heteroscedastic. It is, unfortunately, difficult to avoid this problem entirely. Some options are discussed in a later section of the paper.

Finally, autocorrelation of the residual terms refers to whether there are strings of positive or negative residuals terms. Such strings indicate that the function consistently fits poorly in a certain region of the data, which again suggests that higher order terms, like squares, roots, or interaction terms may need to be added to the function to improve goodness-of-fit. Autocorrelations of residuals were not statistically significant for the two functions developed.

9 PERFORMING SENSITIVITY ANALYSIS

Developing accident modification functions involves a number of analytic choices. Sensitivity analyses should be performed to assess how these choices influence the functions developed. The three most important choices made in developing the accident modification functions for speed enforcement were:

1. First, the decision not to include all the studies that were retrieved, limiting the analysis to studies whose results made sense from a theoretical point of view,

2. Second, the decision to develop two accident modification functions rather than a single function, and

3. Third, the decisions about the mathematical forms of the two functions that were developed.

To test the sensitivity of results to these choices, an analysis was first made including the studies that were excluded from the main analysis, i.e. Andersson (1991), Newstead (2001), Chen et al. (2002), and Goldenbeld and van Schagen (2005). This increased the number of estimates of effect from 27 in the main analysis to 54. Meta-regressions were run on all data points, data points referring to conventional enforcement and data points referring to the use of speed cameras. Estimated model coefficients were compared. The results are presented in Table 6.

Table 6 about here

The left half of the Table contains the results of the main analysis. The right half contains the results of the sensitivity analysis. The results are perfectly consistent as far as the sign of the coefficients is concerned. Not a single coefficient changed sign in the sensitivity analysis when compared to the main analysis. There is, however, a tendency for the coefficients to be attenuated, i.e. be closer to zero and associated with larger standard errors, in the sensitivity analysis. The results of the sensitivity analysis for the use of speed cameras were almost identical to the results of the main

analysis. A single study (Chen et al. 2002) was added in the sensitivity analysis, but it had very low statistical weight and contributed mainly by adding noise to the data.

Based on this analysis, it is concluded that restricting the main analysis to studies with theoretically plausible findings did not fundamentally influence results, but served mainly to reduce noise in the data. Developing two accident modification functions rather than a single one also reduced noise in the data, as evidenced by the fact that the sum of residual variances for the two functions is smaller than the residual variance for the single function fitted to all studies.

It remains to test alternative functional forms. With respect to conventional speed enforcement, the following functional forms were compared:

1. Linear: using effect as dependent variable and level of enforcement as independent variable.

2. Exponential: using the logarithm of effect as dependent variable and level of enforcement as independent variable.

3. Compound: using the logarithm of effect as dependent variable and level of enforcement and level of enforcement squared as independent variables (this was the functional form used in the main analysis).

4. Power: using the logarithm of effect as dependent variable and level of enforcement squared as independent variable.

The inverse function and the logarithmic function could not be tested, as the level of enforcement had the value of zero in one of the studies. As far as studies evaluating the use of speed cameras are concerned, sensitivity analysis tested a linear function, an exponential function, a logarithmic function (the preferred function in the main

analysis) and an inverse function (using 1/camera hours as independent variable). The analyses comparing the different functional firms were made using both the studies selected in the main analysis and the full set of studies included in the sensitivity analysis. The results are reported in table 7.

Table 7 about here

The goodness-of-fit of the various functional forms is indicated by the squared correlation coefficient. The higher the value, the better the fit of the function. It is seen that the chosen functional form fitted better than the alternatives both for studies of conventional speed enforcement and for studies of the use of speed cameras.

10 THE TREATMENT OF HETEROSCEDASTIC DATA

The data used to develop the accident modification functions used for illustrative purposes in this paper were somewhat atypical of data from road safety evaluation studies by not displaying a clear heteroscedasticity. Heteroscedasticity means unequal variance, i.e. that the data points are more widely dispersed in a certain range of outcomes than in another. Figure 7 shows an example of this for studies that have evaluated the effects of bypass roads.

Figure 7 about here

Estimates of effect in studies with large standard errors (i.e. large sampling variance – small accident samples – bottom of diagram) vary much more than those in studies with small standard errors. This means that any function fitted to these data is likely

to fit poorly to the data points characterised by large standard errors. There are four options for dealing with heteroscedasticity:

1. Transform variables to stabilise variance. For variables measured in natural units, a logarithmic transformation will often reduce heteroscedasticity. Experience shows, however, that even after such a transformation considerable heteroscedasticity may remain, as shown in Figure 7.

2. Merge data points, in particular data points that are very widely dispersed, like the data points at the bottom of Figure 7. The drawback of this procedure is that information is lost, for example, when the merged data points referred to different countries or publication years. Moreover, it is difficult to offer definite guidelines on how far one should go in merging data points. Ideally, the practice should be avoided.

3. Restrict the accident modification function to a limited range of the data. This means that the most widely dispersed data points are omitted when developing the accident modification function. Again, this procedure is not ideal since it wastes information.

4. Develop more than one accident modification function. Developing more than one function is relevant when there is reason to believe that the effects of a measure vary in a more complex manner than a single function can represent. Heteroscedasticity as such does not indicate variation in effects and is therefore, by itself, not a sufficient reason to develop more than one function.

None of these options for dealing with heteroscedasticity is ideal. One may have to accept the fact that the residual terms of an accident modification function fitted to heteroscedastic data will, to some extent, be heteroscedastic.

11 UPDATING ACCIDENT MODIFICATION FUNCTIONS

Accident modification functions should be periodically updated. When updating a function, it may be necessary to decide whether to retain the original functional form or change to a new functional form. The following procedure is tentatively proposed for updating an accident modification function.

1. Enter the new data points and estimate predicted values for the entire data set, using both data points that were included before updating and the new data points.

2. If predicted values are close to the original estimates of effect (before adding new data points), the model has been successfully updated.

3. If predicted values fit the original estimates of effect poorer than the original accident modification function, but fit well to the new data points, consider changing functional form.

4. If predicted values fit the original estimates of effect as well as the original accident modification function, but do not fit well to the new estimates of effect, examine more in detail the sources of the poor fit to new estimates of effect. Consider adding variables to the function to improve its fit to the data.

It is stressed that these guidelines are tentative only, as there is so far little experience in updating accident modification functions and few, if any, cases of it have been reported in the literature.

12 DISCUSSION

The case for modelling systematic variation in the effects of by means of accident modification functions has been convincingly made by Hauer et al. (2012). Yet, although it is clear that the effects of very many road safety measures vary systematically, it is by no means trivial to develop functions that describe this variation well and in a useful manner. It is important to emphasise that an accident modification should both be as methodologically rigorous as the data permit (describe variation well) and be applicable when predicting the effects of a measure (be useful).

To develop an accident modification function therefore involves more than a mere historical reconstruction of research. It is an exercise in regression modelling. As such, it faces all the challenges discussed by Hauer (2015) as far as regression modelling in road safety is concerned. Even if everybody can agree on the desirability of developing accident modification functions, this may still often be impossible if, for example:

1. Primary studies employ different designs and are of variable quality.
2. The sources of systematic variation in effects cannot be identified.
3. There are sources of systematic variation in effects that do not produce a systematic pattern in estimates of effects.

It is nearly always the case that road safety evaluation studies differ in design and quality. These differences tend to be associated with systematic variation in estimates of the effects of road safety measures, but not necessarily because the true effects vary. It could, for example, be the case that one study controlled for regression-to-the-mean and the other did not, and that the study not controlling for regression-to-the-mean estimated a larger effect of the road safety measure than the study controlling for regression-to-the-mean.

It is important to remember that systematic variation in effects simply means: "a larger variation than randomness alone can account for". Variation in study quality may well produce systematic variation in this sense, but it is entirely without interest. It is merely a source of confounding, making it more difficult to know whether there really is systematic variation in the effects of road safety measures for reasons that it would be useful for the designers and planners of the measure to know.

One should therefore not be surprised that developing accident modification functions involves many analytic choices. The guidelines presented in this are intended to support these analytic choices. It is hoped that the guidelines are sufficiently clear to give guidance, yet sufficiently "open-ended" not to prevent the exercise of professional judgement, which inevitably will be subjective and frequently open to discussion. It should be clear that developing accident modification functions is, to borrow a term used by Hauer (2015), an "art", not an exact science that can proceed effortlessly by strictly adhering to a well-specified experimental protocol.

13 CONCLUSIONS

Ten methodological guidelines for developing accident modification functions have been proposed and their use illustrated. The guidelines are:

1. Classify, code and select studies that are used as the basis for developing an accident modification function.

2. Perform preparatory analysis to determine the contribution of systematic variation to the overall variation in estimates of effect, to assess the possible presence of publication bias, and to determine whether estimates of effect vary systematically between countries and/or over time.

3. Identify the independent variables of an accident modification function.

4. Identify outlying data points.

5. Identify the best fitting functional form.

6. Determine if more than one function should be develop to describe variation in a data set.

7. Evaluate the quality of the accident modification functions in terms of statistical criteria.

8. Perform sensitivity analysis with respect to analytic choices made as part of the analysis.

9. Assess how best to treat data characterised by heteroscedasticity.

10. Establish a routine for updating the accident modification function.

ACKNOWLEDGEMENT

REFERENCES

Andersson, G. 1991. Effekter på hastigheter av intensifierad övervakning med radar. Transportforskningberedningen (TFB) og Statens väg- och trafikinstitut (VTI) forskning/research nr 6. Stockholm.

Borenstein, M., Hedges, L. V., Higgins, J. T., Rothstein, H. 2009. Introduction to meta-analysis. New York, John Wiley and Sons.

Cameron, M., Newstead, S., Diamantopoulou, K., Oxley, P. 2003. The interaction between speed camera enforcement and speed-related mass media publicity in Victoria. Report 201. Monash University Accident Reseearch Centre, Melbourne.

Chen, G., Meckle, W., Wilson, J. 2002. Speed and safety effect of Photo radar enforcement on a highway corridor in British Columbia. Accident Analysis and Prevention, 34, 129-138.

DeAngelo, G., Hansen, B. 2014. Life and death in the fast lane: Police enforcement and traffic fatalities. American Economic Journal: Economic Policy, 6, 231-257

Duval, S. 2005. The trim and fill method. In Rothstein, H., Sutton, A. J., Borenstein, M. (Eds): Publication bias in meta-analysis: Prevention, assessment and adjustments, 127-144. Chichester, John Wiley and Sons.

Elvik, R. 2011. Developing an accident modification function for speed enforcement. Safety Science, 49, 920-925.

Goldenbeld, C., Schagen, I. van. 2005. The effects of speed enforcement with mobile radar on speed and accidents. An evaluation study on roads in the Dutch province of Friesland. Accident Analysis and Prevention, 37, 1135-1144.

Hauer, E. 2015. The art of regression modelling in road safety. New York, Springer.

Hauer, E., Bamfo, J. 1997. Two tools for finding what function links the dependent variable to the explanatory variables. Proceedings of 10th ICTCT Workshop, Lund, Sweden.

Hauer, E., Bonneson, J., Council, F., Srinivasan, R.., Zegeer, C. 2012. Crash modification factors. Foundational issues. Transportation Research Record, 2279, 67-74.

Lipsey, M. W., Wilson, D. B. 2001. Practical meta-analysis. Thousand Oaks, Sage Publications.

Munden, J. M. 1966. An experiment in enforcing the 30 mile/h speed limit. RRL Report No 24. Harmondsworth, Road Research Laboratory.

Newstead, S. V., Cameron, M. H., Leggett, L. M. W. 2001. The crash reduction effectiveness of a network-wide traffic police deployment system. Accident Analysis and Prevention, 33, 393-406.

Newstead, S., Cameron, M. 2003. Evaluation of the crash effects of the Queensland speed camera program. Report 204. Monash University Accident Research Centre, Melbourne.

Nilsson, E., Engdahl, S. 1982. Effekter av trafikövervakning på trafikolyckor. TOS AB, Stockholm.

Papaioannou, P., Mintsis, G., Taxiltaris. C., Basbas, S. 2002. Enforcement and traffic accidents: recent experience from Greece. Proceedings of 15th ICTCT Workshop, Brno.

Rothstein, H., Sutton, A. J., Borenstein, M. 2005. Publication bias in meta-analysis. Prevention, assessment and adjustments. Chichester, John Wiley and Sons.

Shoup, D. C. 1973. Cost effectiveness of urban traffic law enforcement. Journal of Transport Economics and Policy, 12, 32-57.

Vaa, T., Christensen, P., Ragnøy, A. 1995. Politiets fartskontroller: Virkning på fart og subjektiv oppdagelsesrisiko ved ulike overvåkingsnivåer. TØI rapport 301. Oslo, Transportøkonomisk institutt.

Waard, D. de, Rooijers, T. 1994. An experimental study to evaluate the effectiveness of different methods and intensities of law enforcement on driving speed on motorways. Accident Analysis and Prevention, 26, 751-765.

Yannis, G., Papadimitriou, E., Antoniou, C. 2008. Impact of enforcement on traffic accidents and fatalities: a multivariate, multilevel analysis. Safety Science, 46, 738-750.

LIST OF FIGURES AND TABLES

Figure 1:

Funnel plot of estimates of effect of speed enforcement

Figure 2:

Cumulative residuals plot for function fitted to speed enforcement

Figure 3:

Preliminary function fitted to level of enforcement

Figure 4:

Accident modification function for conventional speed enforcement

Figure 5:

Accident modification function for use of speed cameras

Figure 6:

Graphical test for heteroscedasticity of residuals

Figure 7:

Funnel plot of estimates of effect on accidents of bypass roads

Table 1:

Methodological guidelines for developing accident modification functions

Table 2:

Classification of road safety evaluation studies by design and control for confounding factors

Table 3:

Studies identified on effects of speed enforcement

Table 4:

Data table for studies used to develop accident modification function

Table 5:

Indicators of the quality of accident modification functions

Table 6:

Sensitivity analysis of accident modification functions

Table 7:

Analysis of goodness-of-fit of alternative forms of accident modification functions

Table 1:

Guidelines	Analysis required to comply with guidelines	Justification of guidelines
1. Classify, code and select studies	Classify studies by study design (see Table 2). Do not mix studies employing different designs in the same AMF. Code all variables that may influence effect size.	Studies employing different designs do not control for the same potentially confounding factors. An AMF based on studies employing different designs may be more influenced by confounding than an AMF based on studies employing identical designs.
2. Perform preparatory analysis	The potential presence of publication bias should be tested for. The relative contribution of systematic variation in estimates of effect to overall variance should be quantified. Effects of country and year of publication should be tested for.	An AMF influenced by publication bias will be biased. AMFs should not be developed if publication bias is indicated. An AMF should explain systematic variation in estimates of effect; this only makes sense if systematic variation makes a predominant contribution to the overall variation in estimates of effect. Country and year of publication should be viewed as potentially confounding variables.
3. Identify independent variables	At least one independent variable should be identified. Independent variables may either refer to the measure itself or the context of its use	An AMF should have at least one independent variable. Independent variables should describe characteristics of the measure or the context of its use.
4. Identify outlying data points	Plot data points in a cumulative residuals plot, based on a preliminary AMF, to locate potentially outlying data points. Outlying data points should be omitted.	An outlying data point may decisively influence the mathematical form of an AMF. It is not appropriate that a single data point should determine the shape of a function fitted to, for example, 40-50 data points.
5. Identify the best fitting functional form	A systematic testing of various functional forms, such as linear, power, exponential etc. should be performed in order to identify the best fitting functional form	An AMF can have different functional forms, such as linear, power, exponential, etc. Exploratory testing is needed to identify the best fitting functional form.
6. One or more functions	A careful examination of the residual terms of an AMF can give hints that two or more AMFs are needed to adequately summarise variation in the effects of a measure	The effects of road safety measures may not always be adequately summarised by means of a single AMF. If a more precise description of effects can be obtained by developing more than one AMF, this should be done.
7. Evaluate accident modification function	AMF should be evaluated in terms of predictive performance, explanatory value, and distribution of residual terms	Unless an AMF fits quite well to the data, it cannot be applied to predict the effects of a road safety measure. Several criteria should be applied to assess the quality of an AMF.
8. Perform sensitivity analysis	A sensitivity analysis should be made to assess the effects of analytic choices made when developing an AMF	When developing an AMF analytic choices are made about which studies to include, whether to develop one or more AMFs, the mathematical form of the AMF, and possibly other items. A sensitivity analysis tests how results are influenced by these choices.
9. Decide on treatment of heteroscedasticity	Individual estimates of effect vary in statistical precision. This very often creates unequal variance (heteroscedasticity) across the range covered by the data.	In heteroscedastic data, any function will often fit well to the part of the data characterised by small variance, but poorly to the part of the data characterised by large variance. One should assess options for minimising this problem, although it may be impossible to avoid it entirely.
10. Update accident modification function	A routine for updating AMFs should be established, enabling a decision to made as to whether an updated AMF should retain the original functional form or adopt a new functional form.	AMFs should be periodically updated. When an AMF is updated, rules should be established for either keeping its original mathematical form or changing the mathematical form of the function. If new data points do not fit well to any function, possible reasons for this should be examined.

Table 2:

Main category of study design	Versions of study design by level of control for confounding factors	Rating for study quality (within main group)
Randomised controlled trials (experiments)	Randomised controlled trial demonstrating pre-trial equivalence of groups and controlling for treatment implementation, attrition bias and unintended effects	High
	Randomised controlled trial demonstrating or controlling for some but not all of the factors listed above	Medium
	Randomised controlled trials with evidence of systematic differences between treatment group and control group	Low
Before-and-after studies (observational)	Before-and-after studies controlling for regression-to-the-mean, long-term trends and changes in traffic volume not induced by the measure	High
	Before-and-after studies controlling for some, but not all of the factors listed above	Medium
	Simple before-and-after studies not controlling for any confounding factors	Low
Case-control studies	Case-control studies controlling for self-selection of cases and/or controls and important known risk factors by means of multivariate analysis	High
	Case-control studies controlling partly for self-selection bias and for some but not all known important potentially confounding factors	Medium
	Simple case-control studies not controlling for potentially confounding factors or simple case-series	Low
Cross-sectional studies – multivariate models	Multivariate models not known to be influenced by any of the following potential sources of error: small samples or low mean values; bias due to aggregation or averaging; outlying data points; inclusion of endogenous variables; co-linearity among independent variables; omitted variable bias; wrong functional form; inappropriate model form; inappropriate dependent variable	High
	Multivariate models not known to be influenced by most of the potential sources of error listed above	Medium
	Multivariate models known to be influenced by one or more of the potential sources of error listed above	Low

Table 3:

Study	Study design	Included in original study	Included in current study	Reason for inclusion or exclusion
Munden (1966)	Experimental before-after	Yes	Yes	Internally systematic pattern
Shoup (1973)	Observational before-after	Yes	Yes	Internally systematic pattern
Nilsson and Engdahl (1982)	Observational before-after	Yes	Yes	Internally systematic pattern
Andersson (1991)	Observational before-after	Yes	No	Not internally systematic pattern
Waard and Rooijers (1994)	Experimental before-after	Yes	No	Did not use accidents as dependent variable
Vaa (1995)	Experimental before-after	Yes	No	Did not use accidents as dependent variable
Newstead et al. (2001)	Observational before-after	Yes	No	Not internally systematic pattern
Chen et al. (2002)	Observational before-after	Yes	No	Not internally systematic pattern
Papaioannou et al. (2002)	Observational before-after	Yes	No	Not appropriate dependent variable
Cameron et al. (2003)	Observational before-after	Yes	Yes	Internally systematic pattern
Newstead and Cameron (2003)	Observational before-after	Yes	Yes	Internally systematic pattern
Goldenbeld and van Schagen (2005)	Observational before-after	No	No	Not internally systematic pattern
Yannis et al. (2008)	Observational before-after	No	Yes	Estimate of effect consistent with theoretical prediction
DeAngelo and Hansen (2014)	Observational before-after	No	Yes	Estimate of effect consistent with theoretical prediction

Table 4:

Study	Year	Level of enforcement	Odds ratio	Standard error	Speed camera	Dummy GBR	Dummy USA	Dummy SWE	Dummy AUS	Dummy GRE
Munden	1966	3.5	1.460	0.365	0	1	0	0	0	0
Munden	1966	6.0	0.861	0.285	0	1	0	0	0	0
Munden	1966	8.5	0.815	0.281	0	1	0	0	0	0
Munden	1966	9.0	1.423	0.251	0	1	0	0	0	0
Munden	1966	13.0	0.510	0.297	0	1	0	0	0	0
Munden	1966	14.0	0.637	0.219	0	1	0	0	0	0
Shoup	1973	0.0	1.160	0.193	0	0	1	0	0	0
Shoup	1973	1.0	1.012	0.180	0	0	1	0	0	0
Shoup	1973	1.0	0.870	0.181	0	0	1	0	0	0
Shoup	1973	1.0	1.063	0.187	0	0	1	0	0	0
Shoup	1973	1.1	0.945	0.224	0	0	1	0	0	0
Shoup	1973	3.0	1.056	0.190	0	0	1	0	0	0
Shoup	1973	3.7	0.808	0.213	0	0	1	0	0	0
Shoup	1973	4.0	1.003	0.164	0	0	1	0	0	0
Nilson and Engdahl	1982	0.6	1.091	0.110	0	0	0	1	0	0
Nilson and Engdahl	1982	1.1	0.952	0.080	0	0	0	1	0	0
Nilson and Engdahl	1982	1.8	0.961	0.062	0	0	0	1	0	0
Nilson and Engdahl	1982	3.4	0.880	0.089	0	0	0	1	0	0
Cameron et al.	2003	0.7	1.068	0.042	1	0	0	0	1	0
Cameron et al.	2003	0.9	1.036	0.029	1	0	0	0	1	0
Cameron et al.	2003	1.1	0.979	0.027	1	0	0	0	1	0
Cameron et al.	2003	1.3	0.970	0.028	1	0	0	0	1	0
Newstead and Cameron	2003	1.0	0.938	0.023	1	0	0	0	1	0
Newstead and Cameron	2003	2.3	0.796	0.023	1	0	0	0	1	0
Newstead and Cameron	2003	2.6	0.636	0.022	1	0	0	0	1	0
Newstead and Cameron	2003	3.1	0.610	0.021	1	0	0	0	1	0
Newstead and Cameron	2003	4.4	0.661	0.029	1	0	0	0	1	0
Newstead and Cameron	2003	7.1	0.517	0.018	1	0	0	0	1	0
Yannis et al.	2008	7.9	0.756	0.015	0	0	0	0	0	1
DeAngelo and Hansen	2014	0.7	1.139	0.063	0	0	1	0	0	0
DeAngelo and Hansen	2014	0.7	1.174	0.041	0	0	1	0	0	0

Table 5:

Indicator of model quality	Accident modification functions with respect to	
	Conventional speed enforcement	Speed cameras
Share of systematic variation in primary results (%)	90.3	99.1
Share of systematic variation explained by function (%)	99.4	96.5
Share of cumulative residuals within confidence bounds (%)	82.4	90.0
Unbiasedness of model predictions (predicted/recorded)	1.001	1.005
Normality of standardised residuals	$X^2 = 2.003$; p = 0.849 (normal)	$X^2 = 32.722$; p = 0.000 (not normal)
Heteroscedasticity of standardised residuals	Difference in slope of residuals: –0.123; SE: 0.069, p = 0.062	Difference in slope of residuals: 0.740; SE: 0.143, p = 0.018
Autocorrelation of residuals	Mean for lags 1-15: –0.033; Mean p-value = 0.477	Mean for lags 1-8: –0.063; Mean p-value: 0.172

Table 6:

Models	Main analysis (selected studies)			Sensitivity analysis (all retrieved studies)			
	Term	Estimate	Standard error	Term	Estimate	Standard error	P value for T-test
All studies	Constant	0.1490	0.0529	Constant	0.0448	0.0492	0.1492
	Level of enforcement	-0.0581	0.0104	Level of enforcement	-0.0319	0.0097	0.8968
	Camera dummy	-0.2358	0.0576	Camera dummy	-0.1630	0.0586	0.7654
	Residual variance	0.01481		Residual variance	0.02395		
Conventional	Constant	0.1442	0.0351	Constant	0.0303	0.0527	0.1069
	Level of enforcement	-0.0711	0.0206	Level of enforcement	-0.0376	0.0244	0.7981
	Level squared	0.0022	0.0022	Level squared	0.0011	0.0019	0.3705
	Residual variance	0.00000		Residual variance	0.00832		
Speed cameras	Constant	0.9778	0.0244	Constant	0.9879	0.0264	0.5975
	Level of use (hours)	-0.2540	0.0249	Level of use (hours)	-0.2498	0.0272	0.5401
	Residual variance	0.00249		Residual variance	0.00318		

Table 7:

Functions for conventional speed enforcement (R^2)				Functions for use of speed cameras (R^2)		
Functional form	**Main analysis**	**Sensitivity analysis**	**Functional form**	**Main analysis**	**Sensitivity analysis**
Linear	0.9120	0.1498	Linear	0.7424	0.1565
Exponential	0.9338	0.1833	Exponential	0.7955	0.1787
Compound	0.9398	0.1933	Logarithmic	0.9093	0.4603
Power	0.8675	0.1368	Inverse	0.8860	0.4366

Figure 1:

Funnel plot of estimates of effect of speed enforcement

Axis labels: Fixed-effects standard error (inverted - smallest at top); Weighted mean; Logarithm of estimate of effect

Figure 2:

Cumulative residuals plot for function fitted to level of enforcement

Residuals

Level of enforcement (current level = 1.0)

Preliminary function fitted to level of enforcement

Plot of Accident modification factor (1.00 = no change) versus Level of enforcement (current level = 1.0), ranging from 0.0 to 16.0 on the x-axis and 0.000 to 1.600 on the y-axis.

Legend: Outlying data point; Outlying data point

$y = -0.254\ln(x) + 0.9795$
$R^2 = 0.92$

Legend: ● Data no camera ▲ Data camera ◆ Model no camera ■ Model camera

Figure 4:

Accident modification function for conventional speed enforcement

Figure 5:

Accident modification function for use of speed cameras

Figure 6:

Graphical test for heteroscedasticity of residuals

$y = -0.0808x + 0.8136$
$R^2 = 0.2218$

$y = 0.0418x - 0.8288$
$R^2 = 0.2182$

Standardised residuals

Level of enforcement (1.0 = current level)

Funnel plot of estimates of effect on accidents of bypass roads